An exact bound on the truncated-tilted mean for symmetric distributions

Iosif Pinelis

Department of Mathematical Sciences
Michigan Technological University
Houghton, Michigan 49931, USA
E-mail: ipinelis@mtu.edu

Abstract: An exact upper bound on the Winsorised-tilted mean, $\frac{\mathsf{E}\,X e^{h(X\wedge w)}}{\mathsf{E}\,e^{h(X\wedge w)}}$, of a symmetric random variable X in terms of its second moment is given. Such results are used in work on nonuniform Berry–Esseen-type bounds for general nonlinear statistics.

AMS 2010 subject classifications: Primary 60E15; secondary 60E10, 60F05.
Keywords and phrases: exact upper bounds, Winsorization, truncation, nonuniform Berry-Esseen bounds, Cramér tilt transform, symmetric distributions.

Cramér's tilt transform of a random variable (r.v.) X is a r.v. X_h such that

$$\mathsf{E}\,f(X_h)=\frac{\mathsf{E}\,f(X)e^{h\,X}}{\mathsf{E}\,e^{h\,X}}$$

for all nonnegative Borel functions f, where h is a real parameter. This transform is an important tool in the theory of large deviation probabilities $\mathsf{P}(X>x)$, where $x>0$ is a large number; then the appropriate value of the parameter h is positive. Unfortunately, if the right tail of the distribution of X decreases slower than exponentially, then $\mathsf{E}\,e^{h\,X}=\infty$ for all $h>0$ and thus the tilt transform is not applicable. The usual recourse then is to replace X in the exponent by its truncated counterpart, say $X\,\mathrm{I}\{X\leqslant w\}$ or $X\wedge w$, where w is a real number. As shown in [2, 4], of the two mentioned kinds of truncation, it is the so-called Winsorization, $X\wedge w$, of the r.v. X that is more useful in the applications considered there.

In particular, in [4] one needs a good upper bound on the mean

$$\mathsf{E}_{h,w}\,X:=\frac{\mathsf{E}\,X e^{h(X\wedge w)}}{\mathsf{E}\,e^{h(X\wedge w)}}. \tag{1}$$

of the Winsorised-tilted distribution of X. Note that $\mathsf{E}_{h,w}\,X$ is well defined and finite for any $h\in(0,\infty)$, any $w\in\mathbb{R}$, and any r.v. X such that $\mathsf{E}(0\vee X)<\infty$.

In [2], exact upper bounds on the denominator $\mathsf{E}\,e^{h(X\wedge w)}$ of the ratio in (1) were provided, along with applications to pricing of certain financial derivatives.

Take any positive real numbers h and w. In [1], exact upper bounds on $\mathsf{E}_{h,w} X$ given the first two moments of X. In particular, by [1, Theorem 2.4(II)],

$$\mathsf{E}_{h,w} X < \frac{e^{hw}-1}{w}\, \mathsf{E}\, X^2 \tag{2}$$

for any real-valued r.v. with $\mathsf{E}\, X = 0$ and $\mathsf{E}\, X^2 \in (0,\infty)$; it is also shown in [1] that the factor $\frac{e^{hw}-1}{w}$ in (2) is the best possible one.

The purpose of this note is to show that in the case when (the distribution of) X is symmetric, the factor $\frac{e^{hw}-1}{w}$ in (2) can be improved to $\frac{\operatorname{sh} hw}{w}$; we write sh and ch in place of sinh and cosh.

Theorem 1. *Let X be any* symmetric *real-valued r.v. with $\mathsf{E}\, X^2 \in (0,\infty)$. Then*

$$0 < \mathsf{E}_{h,w} X < \frac{\operatorname{sh} hw}{w}\, \mathsf{E}\, X^2. \tag{3}$$

Remark 2. The factor $\frac{\operatorname{sh} hw}{w}$ in (3) is the best possible one. More specifically,

$$\lim_{\sigma\downarrow 0} \frac{1}{\sigma^2} \sup\big\{ \mathsf{E}_{h,w} X\colon\ \mathsf{E}\, X^2 = \sigma^2,\ X \text{ is symmetric}\big\} = \frac{\operatorname{sh} hw}{w}.$$

In view of Theorem 1, this follows if we let X take values $-w$, 0, and w with probabilities $\frac{\sigma^2}{2w^2}$, $1-\frac{\sigma^2}{w^2}$, and $\frac{\sigma^2}{2w^2}$, respectively, for $\sigma \in (0,w)$, and then let $\sigma \downarrow 0$. Note here that the case of interest in applications in [4] is precisely when $\mathsf{E}\, X^2$ is arbitrarily small. Also, in those applications hw may be rather large, and then the symmetric-case factor $\frac{\operatorname{sh} hw}{w}$ will be almost twice as small as the general zero-mean-case factor $\frac{e^{hw}-1}{w}$.

Proof of Theorem 1. By [1, Proposition 2.6(II)], $\mathsf{E}_{h,w} X$ is increasing in $h>0$, so that $\mathsf{E}_{h,w} X > \mathsf{E}\, X = 0$, and the first inequality in (3) follows.

Let us prove the second inequality in (3). By rescaling, without loss of generality (w.l.o.g.) $h=1$. For all real x and $j \in \{0,1\}$, let

$$f_j(x) := x^j e^{x\wedge w} \quad \text{and} \quad g_j(x) := \tfrac12\big(f_j(x)+f_j(-x)\big),$$

using the convention $0^0 := 1$; then

$$\mathsf{E}\, g_j(|X|) = \mathsf{E}\, f_j(X) = \mathsf{E}\, X^j e^{h(X\wedge w)}. \tag{4}$$

So, (3) will follow if one can show that

$$d := d(u,v,w) := 2\big[g_1(u)+g_1(v) - \tfrac{\operatorname{sh} hw}{w}\, \big(g_0(u)v^2 + g_0(v)u^2\big)\big] < 0 \tag{5}$$

for all positive real u,v,w; indeed, then it will be enough to replace u and v in (5) by independent copies (say U and V) of the r.v. $|X|$, take the expectation, and use (4). At this point, one should note that d may equal 0 if u or v equals 0; in particular, $d=0$ if $u=0$ and $v=w$; however, the condition $\mathsf{E}\, X^2 \in (0,\infty)$

in Theorem 1 implies that $|X| > 0$ with a nonzero probability, which will result in the second inequality in (3) being strict indeed.

So, it remains to prove the inequality (5). Since u and v are interchangeable there, w.l.o.g. $0 < u \leqslant v$. Then (at least) one of the following three cases must occur:

Case 1: $0 < w \leqslant u \leqslant v$;
Case 2: $0 < u \leqslant w \leqslant v$;
Case 3: $0 < u \leqslant v \leqslant w$.

In each of these three cases, d can be expressed without using the minimum function $\wedge$.

In the subsequent treatment of each of these three cases, the default ranges of the variables u, v, and w will be determined by the conditions of the case under consideration. For instance, if in Case 1 (say) an expression in u, v, w is stated to be concave in u or increasing in v, this will mean that it is concave in $u \in [w, v]$ (for any given v and w such that $0 < w \leqslant v$) or, respectively, increasing in $v \in [u, \infty)$ (for any given u and w such that $0 < w \leqslant u$).

As usual, let ∂_z denote the operator of partial differentiation with respect to a variable z.

Case 1

In this case,

$$d = (e^w - e^{-u})u + (e^w - e^{-v})v - \tfrac{\operatorname{sh} w}{w}\big(e^w(u^2 + v^2) + e^{-v}u^2 + e^{-u}v^2\big),$$

whence

$$\partial_v^2 d = e^{-v}(2 - v) - \tfrac{\operatorname{sh} w}{w}(e^{-v}u^2 + 2e^{-u} + 2e^w) \tag{6}$$

and $\partial_v^3 d = e^{-v}(v - 3 + \frac{u^2 \operatorname{sh} w}{w})$. So, $\partial_v^3 d$ may change in sign at most once, and only from $-$ to $+$, if v increases from u to ∞. Therefore,

$$\partial_v^2 d \leqslant (\partial_v^2 d)|_{v=u} \vee (\partial_v^2 d)|_{v=\infty-}. \tag{7}$$

Let $d_2 := d_2(u, w) := we^u(\partial_v^2 d)|_{v=u}$. Then $d_2(u, 0+) = 0$ and $\partial_w d_2 = -2(e^{u+2w}-1) - u - (2+u^2)\operatorname{ch} w < 0$, so that $d_2 < 0$ or, equivalently, $(\partial_v^2 d)|_{v=u} < 0$. It is also clear from (6) that $(\partial_v^2 d)|_{v=\infty-} < 0$. So, by (7), $\partial_v^2 d < 0$ and hence d is strictly concave in $v \in [u, \infty)$.

Therefore, in Case 1 it suffices to show that $d|_{v=u} < 0$ and $(\partial_v d)|_{v=u} < 0$. Introduce $\tilde d := e^{u+w}\,\frac{w}{u}\,d$. Then $\partial_u^2(\tilde d|_{v=u}) = 2e^{u+2w}\big(w - (2 + u)\operatorname{sh} w\big) < 0$, since $\operatorname{sh} w > w$. So, $\tilde d|_{v=u}$ is strictly concave in u. Further, $(\tilde d|_{v=u})|_{u=w} = -(e^w - 1)^3(1 + e^w)w < 0$.

One can see that

$$\big(\partial_u(\tilde d|_{v=u})\big)|_{u=w} = 1 + e^{2w}\big(w + 2e^w w - e^{2w}(1 + w)\big) < 0 \tag{8}$$

for all $w > 0$. Such inequalities, of the form $P(w, e^w) < 0$ for some polynomial P of two variables, can be proved in a rather algorithmic manner. Indeed, let $n \geqslant 1$

be the degree of P in w. "Solving" the inequality $P(w, e^w) < 0$ for w^n, one can rewrite it as $\delta(w) := w^n - P_1(w, e^w) < 0$ or $\delta(w) > 0$ (depending on the sign of the coefficient of w^n in P), where P_1 is some polynomial of degree $\leqslant n-1$ in w. Then $\delta'(w)$ will be a polynomial of degree $\leqslant n-1$ in w, so that one can proceed by induction, ultimately reducing the problem to one on the sign of a polynomial in e^w only. One can use a computer algebra system to (such as Mathematica) to execute such routine calculations, which appears to be a much more reliable and faster way to deal with such matters. In Mathematica, algorithms for solving inequalities like (8) are implemented in the command `Reduce`, which we indeed use to verify (8), as well as a few other similar inequalities. Similar methods were used e.g. in [3].

It follows that $\tilde{d}|_{v=u} < 0$ and hence indeed $d|_{v=u} < 0$. Now (in Case 1) it only remains to verify that $d_1 := d_1(u,w) := we^u(\partial_v d)|_{v=u} < 0$.

Using again the inequality $w < \operatorname{sh} w$ (together with the conditions $0 < w \leqslant u$ of Case 1), one observes that

$$\begin{aligned}\partial_u^2 d_1 &= 2\operatorname{sh} w + e^{u+w}(w - 2(2+u)\operatorname{sh} w) \\ &\leqslant 2\operatorname{sh} w + e^{2w}(w - 2(2+w)\operatorname{sh} w),\end{aligned}$$

and the latter expression can seen to be negative for all $w > 0$ – using again the command `Reduce`, say. So, d_1 is concave in u. Yet another `Reduce` shows that $d_1|_{u=w} < 0$ for $w > 0$. Moreover,

$$(\partial_u d_1)|_{u=w} e^{-w}/2 = (w - \operatorname{sh} w)\operatorname{ch} w - (\operatorname{ch} w + 2w\operatorname{sh} w)\operatorname{sh} w < 0;$$

here we again used the inequality $w < \operatorname{sh} w$. This implies that indeed $d_1 < 0$, which completes the proof of (5) in Case 1.

Case 2

In this case,

$$d = 2u\operatorname{sh} u + v(\operatorname{sh} v + \operatorname{sh} w + \operatorname{ch} w) - v\operatorname{ch} v - \tfrac{\operatorname{sh} w}{w}\big(u^2(e^{-v} + e^w) + 2v^2\operatorname{ch} u\big),$$

whence, introducing

$$d_1 := we^v\,\partial_v d, \tag{9}$$

one has

$$\begin{aligned}e^{-v}\,\partial_v^2\, d_1 &= w\operatorname{ch} w + (w - 4(2+v)\operatorname{ch} u)\operatorname{sh} w \\ &\leqslant w\operatorname{ch} w + (w - 4(2+w))\operatorname{sh} w < 0;\end{aligned}$$

the last inequality here can be obtained via another `Reduce`, and the penultimate inequality follows by the condition $w \leqslant v$ of Case 2. So,

$$d_1 \text{ is concave in } v. \tag{10}$$

Note that the definition (9) of d_1, used in the present Case 2, differs from the definition of d_1 used in Case 1.

Next, $(\partial_v d_1)|_{v=w} = e^w\big(e^w w - 4(w+1)\operatorname{ch} u \operatorname{sh} w\big) + w$ is obviously decreasing in $u > 0$. So, $(\partial_v d_1)|_{v=w} < (\partial_v d_1)|_{v=w,u=0+} = 3w - e^{2w}(w+2) + 2$, which yet another `Reduce` shows to be negative for all $w > 0$. Thus,

$$(\partial_v d_1)|_{v=w} < 0. \tag{11}$$

Now let us show that $d_1|_{v=w} < 0$. One has

$$d_1|_{v=w} = d_{11} + d_{12} \quad \text{and} \quad d_{11} = d_{111} + d_{112}, \tag{12}$$

where

$$\begin{aligned} d_{11} &:= e^w w(\operatorname{sh} w + \operatorname{ch} w - 3\operatorname{ch} u \operatorname{sh} w) + (w-1)w, \\ d_{12} &:= (u^2 - e^w w \operatorname{ch} u)\operatorname{sh} w, \\ d_{111} &:= e^w w(\operatorname{ch} w - 2\operatorname{sh} w) + (w-1)w, \\ d_{112} &:= -3(\operatorname{ch} u - 1)e^w w \operatorname{sh} w. \end{aligned} \tag{13}$$

It is obvious that $d_{112} < 0$. Also, $d_{111} < 0$ by another `Reduce`. Next, $\frac{1}{2\operatorname{sh} w}\partial_u(d_1|_{v=w}) = u - 2e^w w \operatorname{sh} u$. If $u \geqslant 1/2$ then (by the condition $u \leqslant w$ of Case 2) $w \geqslant 1/2$, whence $2e^w w \operatorname{sh} u > \operatorname{sh} u > u$, so that $\partial_u(d_1|_{v=w}) < 0$. Therefore, the condition $\partial_u(d_1|_{v=w}) = 0$ would imply $u < 1/2$ and also $e^w w \operatorname{sh} u = u/2$, and then $u^2 - e^w w \operatorname{ch} u < u^2 - e^w w \operatorname{sh} u = u^2 - u/2 < 0$, so that (by (13)) $d_{12} < 0$ and hence, by (12), $d_1|_{v=w} < 0$. That is, $d_1|_{v=w} < 0$ whenever $\partial_u(d_1|_{v=w}) = 0$.

So, to prove the inequality $d_1|_{v=w} < 0$ it is enough to verify that

$$\begin{aligned} d_1|_{v=w,u=0+} &= (w-1)w + e^w w(\operatorname{ch} w - 3\operatorname{sh} w) < 0 \quad \text{and} \\ \tfrac{1}{w}\, d_1|_{v=w,u=w} &= w + (w + e^w)\operatorname{sh} w - e^w(4\operatorname{sh} w - 1)\operatorname{ch} w - 1 < 0, \end{aligned}$$

which again can be done using `Reduce`. We conclude that indeed $d_1|_{v=w} < 0$.

Using also the earlier established conditions (10) and (11), as well as the Case 2 condition $v \geqslant w$, one has $d_1 < 0$. So, by (9), d is decreasing in v.

To complete the consideration of Case 2, it remains to show that $d|_{v=w} < 0$. Observe here that

$$\begin{aligned} \tfrac{1}{2}\,\partial_u(d|_{v=w}) &= \operatorname{sh} u + u\operatorname{ch} u - 2\,\tfrac{\operatorname{sh} w}{w}\, u \operatorname{ch} w - w \operatorname{sh} u \operatorname{sh} w \\ &< \operatorname{sh} u + u\operatorname{ch} u - 2u\operatorname{ch} w \leqslant \operatorname{sh} u - u \operatorname{ch} u < 0, \end{aligned}$$

so that $d|_{v=w}$ is decreasing in $u > 0$, whereas $d|_{v=w,u=0+} = 0$. Thus, indeed $d|_{v=w} < 0$, and (5) is proved in Case 2 as well.

It remains to consider

Case 3

Note that $\frac{\operatorname{sh} w}{w}$ is increasing in $w > 0$. So, by (5), d is decreasing in $w \in [v, \infty)$, because $g_j(u)$ and $g_j(v)$ do not depend on w as long as $w \geqslant u \vee v$. It follows that in Case 3 w.l.o.g. $w = v$. Thus, $0 < u \leqslant w = v$, so that Case 3 has been quickly reduced to the already considered Case 2.

Now inequality (5) and thereby Theorem 1 are completely proved. □

References

[1] Pinelis, I. Exact bounds on the truncated-tilted mean, with applications (preprint). http://arxiv.org/abs/1103.3683, 2011.

[2] Pinelis, I. Exact lower bounds on the exponential moments of Winsorized and truncated random variables. *J. App. Probab. 48* (2011), 547–560.

[3] Pinelis, I. An asymptotically Gaussian bound on the Rademacher tails. *Electron. J. Probab. 17* (2012), 1–22.

[4] Pinelis, I., and Molzon, R. Berry-Esséen bounds for general nonlinear statistics, with applications to Pearson's and non-central Student's and Hotelling's (preprint), arXiv:0906.0177v1 [math.ST].